

ON MIXING PROPERTIES OF SOME INAR MODELS

RICHARD C. BRADLEY

Dedicated to the memory of Mikhail Gordin

ABSTRACT. Strictly stationary INAR(1) (“integer-valued autoregressive processes of order 1”) with Poisson innovations are “interlaced ρ -mixing”.

1. INTRODUCTION

The INAR processes, or “integer-valued autoregressive processes”, are a variant of the usual autoregressive processes in time series analysis. In various references, the INAR processes and variations on them have been studied as models to use in the statistical analysis of “count data”. See e.g. [8], [12], [14], [17], [18], and the references therein. In [17], for certain INAR(1) processes (“integer-valued autoregressive processes of order 1”), and certain variations on them, strong mixing and even absolute regularity were verified, with exponential mixing rate.

In the study of INAR processes, the ρ -mixing condition does not seem to have gotten much attention, but it could perhaps play a useful role as well, given the extensive literature on limit theory under ρ -mixing that has been developed since early results on that topic such as in [10] and [16].

This note here will go in a little different direction. Within the INAR processes, one particularly prominent subclass is the strictly stationary INAR(1) processes with “Poisson innovations”. For that subclass (and some other related processes), absolute regularity with exponential mixing rate was already verified in [17]. In this note, for that prominent subclass, we shall verify the ρ^* -mixing (“interlaced ρ -mixing”) condition, which is stronger than ρ -mixing. (Both of those latter two mixing conditions and that subclass of processes will be explicitly formulated below.) The Poisson innovations seem to facilitate the study of the ρ^* -mixing condition for that subclass. The techniques in this note involving ρ^* -mixing can apparently be extended to some limited extent to some other INAR processes, and even to some variations on them such as ones in [17]. For simplicity, this note will be confined to just the subclass identified above.

The processes in the prominent subclass discussed above are strictly stationary, countable-state Markov chains. It is well known and elementary that for Markov chains, for either the ρ -mixing condition or the ρ^* -mixing condition, the mixing rate is automatically (at least) exponential. Now strictly stationary, finite-state, irreducible, aperiodic Markov chains are ρ^* -mixing (see [1] or [3, Theorem 7.15]). However, for strictly stationary, countable-state Markov chains in general, ρ -mixing does not imply ρ^* -mixing. (Counterexamples are constructed in [2] and [4], with the ones in the latter reference being reversible.) For the INAR processes in general, and in particular for the (Markovian) INAR(1) processes whose innovations are not

Poisson, there is more to explore regarding the ρ -mixing and ρ^* -mixing conditions and the connections between them.

Now let us formulate the ρ -mixing and ρ^* -mixing conditions, define the subclass of processes that will be studied here, and then give the main result.

Suppose $X := (X_k, k \in \mathbb{Z})$ is a strictly stationary sequence of random variables on a probability space (Ω, \mathcal{F}, P) . For any two σ -fields \mathcal{A} and $\mathcal{B} \subset \mathcal{F}$, define the “maximal correlation” [9]:

$$\rho(\mathcal{A}, \mathcal{B}) := \sup |\text{Corr}(f, g)|$$

where the supremum is taken over all pairs of square-integrable random variables f and g such that f is \mathcal{A} -measurable and g is \mathcal{B} -measurable. For each positive integer n , define the following two dependence coefficients (for the given strictly stationary sequence X):

$$\rho(X, n) := \rho(\sigma(X_k, k \leq 0), \sigma(X_k, k \geq n)) \quad (1.1)$$

and

$$\rho^*(X, n) := \sup \rho(\sigma(X_k, k \in S), \sigma(X_k, k \in T)) \quad (1.2)$$

where the supremum is taken over all pairs of nonempty, disjoint sets $S, T \subset \mathbb{Z}$ such that

$$\text{dist}(S, T) := \min_{s \in S, t \in T} |s - t| \geq n. \quad (1.3)$$

In (1.1), (1.2), and below, the notation $\sigma(\dots)$ means the σ -field generated by (\dots) . In (1.2)–(1.3), the sets S and T can be “interlaced”, with each set containing elements between ones in the other set. The (strictly stationary) sequence X is said to be “ ρ -mixing” (a condition introduced in [11]) if $\rho(X, n) \rightarrow 0$ as $n \rightarrow \infty$, and ρ^* -mixing (a condition apparently first studied in [19]) if $\rho^*(X, n) \rightarrow 0$ as $n \rightarrow \infty$. Obviously $\rho(X, n) \leq \rho^*(X, n)$ for each $n \geq 1$, and (hence) ρ^* -mixing implies ρ -mixing.

The following terminology will be useful.

Definition 1.1. *An ordered triplet $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ of σ -fields ($\subset \mathcal{F}$) will be called a “Markov triplet” if any (hence all) of the following three equivalent conditions holds:*

- (i) *for all $C \in \mathcal{C}$, $P(C | \mathcal{A} \vee \mathcal{B}) = P(C | \mathcal{B})$ a.s.;*
- (ii) *for all $A \in \mathcal{A}$ and all $C \in \mathcal{C}$, $P(A \cap C | \mathcal{B}) = P(A | \mathcal{B}) \cdot P(C | \mathcal{B})$ a.s.;*
- (iii) *for all $A \in \mathcal{A}$, $P(A | \mathcal{B} \vee \mathcal{C}) = P(A | \mathcal{B})$ a.s.*

The following elementary observation will be useful later on: If $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ is a Markov triplet, then $(\mathcal{A} \vee \mathcal{B}, \mathcal{B}, \mathcal{B} \vee \mathcal{C})$ is a Markov triplet, and (hence) for any σ -fields $\mathcal{G} \subset \mathcal{A} \vee \mathcal{B}$ and $\mathcal{H} \subset \mathcal{B} \vee \mathcal{C}$, $(\mathcal{G}, \mathcal{B}, \mathcal{H})$ is a Markov triplet.

In what follows, \mathbb{N} denotes the set of all positive integers, and $\overline{\mathbb{N}} := \mathbb{N} \cup \{0\}$ denotes the set of all nonnegative integers.

Definition 1.2. *Suppose $a \in (0, 1)$ and $\lambda > 0$. A strictly stationary “INAR(1) process with Poisson innovations” (with parameters a and λ), is a strictly stationary Markov chain $X := (X_k, k \in \mathbb{Z})$ with state space $\overline{\mathbb{N}}$, with X having the following “structural” properties: There exist random variables U_k, V_k , $k \in \mathbb{Z}$ for which the following conditions hold:*

- (i) *For each $k \in \mathbb{Z}$, $X_k = U_k + V_k$.*
- (ii) *For each $k \in \mathbb{Z}$ and each $x \in \overline{\mathbb{N}}$, the conditional distribution of U_k given $\{X_{k-1} = x\}$ is binomial with parameters x and a .*

(iii) For each $k \in \mathbb{Z}$, the ordered triplet of σ -fields

$$(\sigma(U_j, V_j, X_j, j \leq k-1), \sigma(X_{k-1}), \sigma(U_k))$$

is a Markov triplet.

(iv) For each $k \in \mathbb{Z}$, the distribution of the random variable V_k is Poisson with mean λ .

(v) For each $k \in \mathbb{Z}$, the random variable V_k is independent of the σ -field $\sigma(U_j, V_j, X_j, j \leq k-1) \vee \sigma(U_k)$.

In Definition 1.2, for a given $k \in \mathbb{Z}$, the random variable V_k is the “Poisson innovation”. It is well known and elementary (see e.g. [17]) that in the context of Definition 1.2, the (invariant) marginal distribution of each X_k is Poisson with mean $\lambda/(1-a)$.

Here is the main result of this note:

Theorem 1.3. *Suppose $a \in (0, 1)$ and $\lambda > 0$; and suppose $X := (X_k, k \in \mathbb{Z})$ is the strictly stationary INAR(1) process (Markov chain) in Definition 1.2, meeting all conditions there (including the Poisson (λ) “innovations”). Then X is ρ^* -mixing (with $\rho^*(X, n) \rightarrow 0$ at least exponentially fast as $n \rightarrow \infty$).*

The proof of this theorem will be carried out through Sections 2, 3, and 4 below. From that proof, one can see that the (of course exponential) mixing rate for ρ^* -mixing in Theorem 1.3 essentially depends only on (an upper bound for) the parameter a , not on λ .

2. PRELIMINARIES

Throughout the rest of this note, the setting will be a probability space (Ω, \mathcal{F}, P) , rich enough to accommodate all random variables specified. Random variables are real-valued (and often integer-valued or even $\{0, 1\}$ -valued) unless specified otherwise.

Section 2 here will be devoted to some lemmas that will be used in the proof of Theorem 1.3.

The following lemma is due to Csáki and Fisher [7]. (The proof given there has a flaw. For a fully correct proof, see [20] or [3, Theorem 6.1].)

Lemma 2.1. *Suppose \mathcal{A}_n and \mathcal{B}_n , $n \in \mathbb{N}$ are σ -fields ($\subset \mathcal{F}$), and the σ -fields $\mathcal{A}_n \vee \mathcal{B}_n$, $n \in \mathbb{N}$ are independent. Then*

$$\rho \left(\bigvee_{n \in \mathbb{N}} \mathcal{A}_n, \bigvee_{n \in \mathbb{N}} \mathcal{B}_n \right) = \sup_{n \in \mathbb{N}} \rho(\mathcal{A}_n, \mathcal{B}_n).$$

Next, for any two σ -fields \mathcal{A} and \mathcal{B} ($\subset \mathcal{F}$), define the following measure of dependence:

$$\lambda(\mathcal{A}, \mathcal{B}) := \sup \frac{|P(A \cap B) - P(A)P(B)|}{[P(A)]^{1/2}[P(B)]^{1/2}} \quad (2.1)$$

where the supremum is taken over all pairs of events $A \in \mathcal{A}$ and $B \in \mathcal{B}$ such that $P(A) > 0$ and $P(B) > 0$.

Lemma 2.2. *For any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that the following holds: If \mathcal{A} and \mathcal{B} are σ -fields such that $\lambda(\mathcal{A}, \mathcal{B}) \leq \delta$, then $\rho(\mathcal{A}, \mathcal{B}) \leq \varepsilon$.*

Quite sharp versions of Lemma 2.2 can be found in [5], [6], [3, Theorem 4.15], and in a very sharp form, [15].

Lemma 2.3. *Suppose $0 < \varepsilon \leq 1/9$. Suppose (X_1, X_2, X_3, \dots) is a sequence of random variables such that for each $n \geq 2$, $P(X_n = 0 \mid X_{n-1} = 0) = 1$ and*

$$P(X_n = 0 \mid \sigma(X_1, X_2, \dots, X_{n-1})) \geq 1 - \varepsilon \text{ a.s.}$$

Then

$$\lambda(\sigma(X_1, X_3, X_5, X_7, \dots), \sigma(X_2, X_4, X_6, X_8, \dots)) \leq 3\varepsilon^{1/2}.$$

A proof of Lemma 2.3 can be found in [1, Lemma 3.1] or [3, Theorem 5.21]. (In Lemma 2.3, the labeling of the “absorbing state” as 0 is just for convenience.)

Next, for any (not necessarily stationary) sequence $X := (X_k, k \in \mathbb{Z})$ or $X := (X_k, k \in \overline{\mathbb{N}})$, define the dependence coefficients $\rho^*(n)$, $n \in \mathbb{N}$ by (1.2)–(1.3). (In the case of index set $\overline{\mathbb{N}}$, the sets S and T are restricted to that set.)

In what follows, if S is a nonempty finite set $\subset \overline{\mathbb{N}}$, J is its cardinality, X_k , $k \in S$ are random variables, and (say) $f : \mathbb{R}^J \rightarrow \mathbb{R}$ is a Borel function, then the notation $f(X_k, k \in S)$ means $f(X_{k(1)}, X_{k(2)}, \dots, X_{k(J)})$ where $k(1) < k(2) < \dots < k(J)$ are the elements of S in strictly increasing order.

Lemma 2.4. *For any $a \in (0, 1)$ and any $\varepsilon > 0$, there exists a positive integer $m = m(a, \varepsilon)$ such that the following holds:*

Suppose ζ_0 is a $\{0, 1\}$ -valued random variable. Suppose $\eta := (\eta_1, \eta_2, \eta_3, \dots)$ is a sequence of independent, identically distributed $\{0, 1\}$ -valued random variables such that $P(\eta_1 = 1) = a$, with this sequence η being independent of ζ_0 . For each $k \in \mathbb{N}$, define the $\{0, 1\}$ -valued random variable

$$\zeta_k := \zeta_0 \cdot \prod_{i=1}^k \eta_i. \quad (2.2)$$

Then the random sequence $\zeta := (\zeta_0, \zeta_1, \zeta_2, \dots)$ satisfies

$$\rho^*(\zeta, m) \leq \varepsilon. \quad (2.3)$$

Proof. Suppose $a \in (0, 1)$ and $\varepsilon > 0$. Our first task is to define the positive integer $m = m(a, \varepsilon)$.

Referring to (2.1), let $\delta = \delta(\varepsilon) > 0$ be as in Lemma 2.2. Let $\gamma \in (0, 1/9]$ be such that

$$3\gamma^{1/2} \leq \delta. \quad (2.4)$$

Note that δ and (hence) γ depend only on ε . Let $m = m(a, \varepsilon)$ be a positive integer such that

$$a^m \leq \gamma. \quad (2.5)$$

That completes the definition of $m = m(a, \varepsilon)$.

Now suppose the random variable ζ_0 , the random sequence η , and (then) the random sequence ζ are as in the statement of Lemma 2.4. Our task is to prove (2.3).

Suppose S and T are any two nonempty, disjoint subsets of $\overline{\mathbb{N}}$ such that $\text{dist}(S, T) \geq m$. To complete the proof of (2.3), it suffices to show that

$$\rho(\sigma(\zeta_k, k \in S), \sigma(\zeta_k, k \in T)) \leq \varepsilon. \quad (2.6)$$

By a standard measure-theoretic argument, it suffices to show (2.6) in the case where both index sets S and T are finite. We make that assumption.

Just for convenience, without loss of generality (after switching S and T if necessary, and after enlarging T by one element if necessary), we assume that

the least and greatest elements of the set $S \cup T$ belong to S and T respectively. Then there exists a positive even integer L and nonempty, (pairwise) disjoint sets $Q_1, Q_2, \dots, Q_L \subset \overline{\mathbb{N}}$ with the following properties:

$$\begin{aligned} S &= \bigcup_{i \in \{1, 3, 5, \dots, L-1\}} Q_i; \\ T &= \bigcup_{i \in \{2, 4, 6, \dots, L\}} Q_i; \text{ and} \\ \forall i \in \{1, 2, \dots, L-1\}, \quad m + [\max Q_i] &\leq [\min Q_{i+1}]. \end{aligned} \quad (2.7)$$

For each positive integer J , let $\phi_J : \{0, 1\}^J \rightarrow \overline{\mathbb{N}}$ be a one-to-one function such that $\phi_J(0, 0, \dots, 0) = 0$. For each $i \in \{1, 2, \dots, L\}$, define the (\mathbb{N} -valued) random variable $X_i = \phi_{J(i)}(\zeta_k, k \in Q_i)$ where $J(i)$ is the cardinality of Q_i . Then

$$\begin{aligned} \forall i \in \{1, 2, \dots, L\}, \\ \sigma(X_i) = \sigma(\zeta_k, k \in Q_i) \text{ and } \{X_i = 0\} &= \{\zeta_k = 0 \ \forall k \in Q_i\}; \end{aligned} \quad (2.8)$$

and (hence)

$$\begin{aligned} \sigma(\zeta_k, k \in S) &= \sigma(X_1, X_3, X_5, \dots, X_{L-1}) \text{ and} \\ \sigma(\zeta_k, k \in T) &= \sigma(X_2, X_4, X_6, \dots, X_L). \end{aligned} \quad (2.9)$$

For each $k \in \mathbb{N}$, by (2.2) and the assumptions in Lemma 2.4, one has that (i) $\zeta_k = \zeta_{k-1} \cdot \eta_k$ and hence $\{\zeta_{k-1} = 0\} \subset \{\zeta_k = 0\}$, and (ii) the σ -fields $\sigma(\eta_i, i \geq k)$ and $\sigma(\zeta_i, i \leq k-1)$ are independent. These facts have the following two consequences:

First, by (2.7) and (2.8), for each $i \in \{2, 3, \dots, L\}$, $\{X_{i-1} = 0\} \subset \{X_i = 0\}$ and hence $P(X_i = 0 \mid X_{i-1} = 0) = 1$.

Second, for each $i \in \{2, 3, \dots, L\}$, letting $j := \max Q_{i-1}$, one has by (2.2), (2.7), and (2.8) that $\{X_i = 0\} \supset \bigcup_{u=j+1}^{j+m} \{\eta_u = 0\}$, this latter event is independent of $\sigma(\zeta_k, k \leq j)$ and hence independent of $\sigma(X_1, X_2, \dots, X_{i-1})$, and hence now by (2.5), almost surely

$$\begin{aligned} P(X_i = 0 \mid \sigma(X_1, X_2, \dots, X_{i-1})) &\geq P\left(\bigcup_{u=j+1}^{j+m} \{\eta_u = 0\} \mid \sigma(X_1, X_2, \dots, X_{i-1})\right) \\ &= P\left(\bigcup_{u=j+1}^{j+m} \{\eta_u = 0\}\right) \\ &= 1 - P\left(\bigcap_{u=j+1}^{j+m} \{\eta_u = 1\}\right) \\ &= 1 - a^m \geq 1 - \gamma. \end{aligned}$$

It now follows from (2.9), Lemma 2.3, and (2.4) that

$$\begin{aligned} \lambda(\sigma(\zeta_k, k \in S), \sigma(\zeta_k, k \in T)) &= \\ &= \lambda(\sigma(X_1, X_3, X_5, \dots, X_{L-1}), \sigma(X_2, X_4, X_6, \dots, X_L)) \leq 3\gamma^{1/2} \leq \delta. \end{aligned}$$

Hence by the definition of δ (just before (2.4), and based on Lemma 2.2), (2.6) holds. That completes the proof. \square

Note that by adapting the proof of Lemma 2.4, one can extend Lemma 2.4 to the broader class of random sequences in the hypothesis of Lemma 2.3, with the $\varepsilon \leq 1/9$ there replaced by $a \in (0, 1)$. However, Lemma 2.4 in its present form will suffice for our purposes here.

This section will conclude with a lemma giving just a few related standard elementary facts which will be used later on. Here and below, for a given $a \in (0, 1)$, the “binomial distribution with parameters 0 and a ” is of course the point mass at 0.

Lemma 2.5. *Suppose $a \in (0, 1)$. Suppose $\lambda_1, \lambda_2, \lambda_3 \dots$ is a sequence of positive numbers such that $\sum_{i=1}^{\infty} \lambda_i < \infty$. Suppose $(Y_1, Z_1), (Y_2, Z_2), (Y_3, Z_3), \dots$ is a sequence of independent random vectors such that for each $i \in \mathbb{N}$, (i) the distribution of Y_i is Poisson with mean λ_i , and (ii) for each $y \in \overline{\mathbb{N}}$, the conditional distribution of Z_i given $\{Y_i = y\}$ is binomial with parameters y and a .*

(A) *Then $Y := \sum_{i=1}^{\infty} Y_i < \infty$ a.s., and this random variable Y has the Poisson distribution with mean $\sum_{i=1}^{\infty} \lambda_i$.*

(B) *Also, $Z := \sum_{i=1}^{\infty} Z_i \leq Y < \infty$ a.s. Further, for any $y \in \overline{\mathbb{N}}$, the conditional distribution of Z given $\{Y = y\}$ is binomial with parameters y and a .*

(C) *The ordered triplet of σ -fields $(\sigma(Y_i, i \in \mathbb{N}), \sigma(Y), \sigma(Z))$ is a Markov triplet.*

Statement (A) holds by a simple limiting argument. Statements (B) and (C) both follow from the elementary fact that if m is a nonnegative integer and (y_1, y_2, y_3, \dots) is a sequence of nonnegative integers whose sum is m (which allows at most finitely many y_i 's to be nonzero), then the event $\bigcap_{i=1}^{\infty} \{Y_i = y_i\}$ has positive probability and is an atom of the σ -field $\sigma(Y_1, Y_2, Y_3, \dots)$, and the conditional distribution of Z given that event is binomial with parameters m and a .

3. TWO MARKOV CHAINS

In this section, in preparation for the main argument for Theorem 1.3 to be given in Section 4, the property of ρ^* -mixing will be verified for two classes of (nonstationary) Markov chains.

Lemma 3.1. *Suppose $a \in (0, 1)$, $p \in (0, 1)$, and $N \in \mathbb{N}$. Suppose $Y := (Y_0, Y_1, Y_2, \dots)$ is a Markov chain whose states are nonnegative integers, such that (i) the distribution of Y_0 is binomial (N, p) , and (ii) for each $j \in \overline{\mathbb{N}}$ and each integer y such that $P(Y_j = y) > 0$, the conditional distribution of Y_{j+1} given $\{Y_j = y\}$ is binomial (y, a) .*

Suppose $\varepsilon > 0$, and the positive integer $m = m(a, \varepsilon)$ is as in Lemma 2.4. Then

$$\rho^*(Y, m) \leq \varepsilon. \quad (3.1)$$

Proof. By a standard measure-theoretic argument, the dependence coefficients $\rho^*(\cdot, n)$, $n \in \mathbb{N}$ for a given random sequence depend only on the distribution of that whole random sequence. Also, the distribution of a (say discrete-state) Markov chain $Y := (Y_0, Y_1, Y_2, \dots)$ is uniquely determined by the marginal distribution of Y_0 and the one-step transition probabilities. Hence it suffices to carry out the proof of Lemma 3.1 for a Markov chain Y that satisfies the conditions in Lemma 3.1 and is embedded in a convenient context.

Refer to the parameters a , p , and N in the statement of Lemma 3.1. Let $\eta := (\eta_{h,j}, 1 \leq h \leq N, j \in \mathbb{N})$ be an array of independent, identically distributed $\{0, 1\}$ -valued random variables such that $\rho(\eta_{1,1} = 1) = a$.

Let $\zeta := (\zeta_{h,j}, 1 \leq h \leq N, j \in \overline{\mathbb{N}})$ be an array of $\{0, 1\}$ -valued random variables that meets the following two conditions (interpreted appropriately if $N = 1$): (i) The random variables $\zeta_{h,0}, 1 \leq h \leq N$ are independent, identically distributed $\{0, 1\}$ -valued random variables such that $P(\zeta_{1,0} = 1) = p$, with the sequence $(\zeta_{h,0}, 1 \leq h \leq N)$ being independent of the array η . (ii) For each $h \in \{1, 2, \dots, N\}$ and each $j \in \mathbb{N}$,

$$\zeta_{h,j} := \zeta_{h,0} \cdot \prod_{i=1}^j \eta_{h,i}. \quad (3.2)$$

Define the sequence $Y := (Y_0, Y_1, Y_2, \dots)$ of (nonnegative, integer-valued) random variables as follows: For each $j \in \mathbb{N}$,

$$Y_j := \sum_{h=1}^N \zeta_{h,j}. \quad (3.3)$$

By (3.2), for every $h \in \{1, 2, \dots, N\}$ and every $j \in \overline{\mathbb{N}}$,

$$\zeta_{h,j+1} = \zeta_{h,j} \cdot \eta_{h,j+1}. \quad (3.4)$$

By (3.3) and (3.4),

$$N \geq Y_0 \geq Y_1 \geq Y_2 \geq \dots \geq 0. \quad (3.5)$$

By (3.3) and the properties of the array ζ ,

$$\text{the distribution of } Y_0 \text{ is binomial } (N, p). \quad (3.6)$$

Our next task, starting with (3.6), is to establish the distribution of the entire sequence Y .

Define (with some redundancy) the σ -fields $\mathcal{G}_j, j \in \overline{\mathbb{N}}$ as follows:

$$\begin{aligned} \mathcal{G}_0 &:= \sigma(\zeta_{h,0}, 1 \leq h \leq N); \text{ and} \\ \forall j \in \mathbb{N}, \mathcal{G}_j &:= \sigma(\zeta_{h,k}, 1 \leq h \leq N, 0 \leq k \leq j) \vee \sigma(\eta_{h,k}, 1 \leq h \leq N, 1 \leq k \leq j). \end{aligned} \quad (3.7)$$

For each $j \in \overline{\mathbb{N}}$, the σ -field \mathcal{G}_j is independent of $\sigma(\eta_{h,k}, 1 \leq h \leq N, k \geq j+1)$.

Now suppose $j \in \overline{\mathbb{N}}$; and suppose $y \in \{1, 2, \dots, N\}$, and $S \subset \{1, 2, \dots, N\}$ is a set with cardinality y . Define the event

$$A := \{\forall h \in S, \zeta_{h,j} = 1; \text{ and } \forall h \in \{1, \dots, N\} - S, \zeta_{h,j} = 0\}. \quad (3.8)$$

(If $y = N$ then $A = \bigcap_{h=1}^N \{\zeta_{h,j} = 1\}$.) Suppose $G \in \mathcal{G}_j$ (see (3.7)) is an event, and that $P(G \cap A) > 0$. Then $Y_{j+1} = \sum_{h=1}^N \zeta_{h,j} \cdot \eta_{h,j+1}$ by (3.3) and (3.4); and hence by the sentence after (3.7) and a simple argument, for every $z \in \{0, 1, \dots, y\}$,

$$P(Y_{j+1} = z \mid G \cap A) = \binom{y}{z} a^z (1-a)^{y-z}. \quad (3.9)$$

Next suppose again that $j \in \overline{\mathbb{N}}$ and $y \in \{1, 2, \dots, N\}$. By (3.3), the event $\{Y_j = y\}$ is the union of finitely many (pairwise) disjoint events of the form A in (3.8). Hence by (3.9) and a simple calculation, if $G \in \mathcal{G}_j$, $P(G \cap \{Y_j = y\}) > 0$, and $z \in \{0, 1, \dots, y\}$, then

$$P(Y_{j+1} = z \mid G \cap \{Y_j = y\}) = \binom{y}{z} a^z (1-a)^{y-z}. \quad (3.10)$$

Of course (recall (3.5)) eq. (3.10) also holds for $y = 0$ (and $z = 0$). Also, by (3.3) and (3.7), each of the random variables Y_k , $0 \leq k \leq j$ is \mathcal{G}_j -measurable. Hence (3.10) has the following consequences:

The sequence Y is a Markov chain. For every $j \in \overline{\mathbb{N}}$ and every $y \in \{0, 1, \dots, N\}$, $P(Y_j = y) > 0$ (by (3.6) followed by (3.10) and induction, with $G = \Omega$). Finally, for each $j \in \overline{\mathbb{N}}$ and each $y \in \{0, 1, \dots, N\}$, the conditional distribution of Y_{j+1} given $\{Y_j = y\}$ is binomial (y, a) . Hence by (3.6), the sequence Y meets all conditions specified in Lemma 3.1.

Now suppose $\varepsilon > 0$, and $m = m(a, \varepsilon)$ is as in Lemma 2.4. To complete the proof of Lemma 3.1, it suffices to prove for the sequence Y above that (3.1) holds.

For each $h \in \{1, 2, \dots, N\}$, define the random sequence $\zeta^{(h)} := (\zeta_{h,0}, \zeta_{h,1}, \zeta_{h,2}, \dots)$. By (3.2) and the properties of the arrays η and ζ here, for each $h \in \{1, 2, \dots, N\}$, the sequence $\zeta^{(h)}$ fulfills the conditions in Lemma 2.4. Hence from Lemma 2.4,

$$\forall h \in \{1, \dots, N\}, \quad \rho^*(\zeta^{(h)}, m) \leq \varepsilon. \quad (3.11)$$

Also, by (3.2) and the properties of the arrays η and ζ here, the sequences $\zeta^{(h)}$, $h \in \{1, 2, \dots, N\}$ are independent of each other. Hence by (3.3), (3.11), and Lemma 2.1, eq. (3.1) holds. That completes the proof. \square

Lemma 3.2. *Suppose $a \in (0, 1)$ and $\lambda > 0$. Suppose $Y := (Y_0, Y_1, Y_2, \dots)$ is a Markov chain with state space $\overline{\mathbb{N}}$, such that (i) the distribution of the random variable Y_0 is Poisson (λ) , and (ii) for each $j \in \overline{\mathbb{N}}$ and each $y \in \overline{\mathbb{N}}$, the conditional distribution of Y_{j+1} given $\{Y_j = y\}$ is binomial (y, a) .*

(A) *For each $j \in \overline{\mathbb{N}}$, the distribution of the random variable Y_j is Poisson (λa^j) .*

(B) *Suppose $\varepsilon > 0$, and suppose the positive integer $m = m(a, \varepsilon)$ is as in Lemma 2.4. Then $\rho^*(Y, m) \leq \varepsilon$.*

Proof. For statement (A), conditions (i) and (ii) in Lemma 3.2 imply that Y_1 is Poisson (λa) by a standard calculation, and by repeating that argument one obtains (A) by induction.

Proof of (B). For each integer $n > \lambda$, let $Y^{(n)} := (Y_0^{(n)}, Y_1^{(n)}, Y_2^{(n)}, \dots)$ be a Markov chain with state space $\{0, 1, \dots, n\}$ such that (i) the distribution of $Y_0^{(n)}$ is binomial $(n, \lambda/n)$, and (ii) for each $j \in \overline{\mathbb{N}}$ and each $y \in \{0, 1, \dots, n\}$, the conditional distribution of $Y_{j+1}^{(n)}$ given $\{Y_j = y\}$ is binomial (y, a) . Then $Y_0^{(n)}$ converges in distribution to Y_0 (which is Poisson (λ)) as $n \rightarrow \infty$. Since the one-step transition probabilities for each of the Markov chains $Y^{(n)}$ are the same as for the Markov chain Y , one has that for every $j \in \overline{\mathbb{N}}$ and every choice of nonnegative integers y_0, y_1, \dots, y_j ,

$$P\left(\bigcap_{i=0}^j \{Y_i^{(n)} = y_i\}\right) \longrightarrow P\left(\bigcap_{i=0}^j \{Y_i = y_i\}\right) \text{ as } n \rightarrow \infty. \quad (3.12)$$

The rest of this argument is routine, but let us go through it. Suppose $\varepsilon > 0$, and suppose $m = m(a, \varepsilon)$ is as in Lemma 2.4. Suppose S and T are nonempty, finite, disjoint subsets of $\overline{\mathbb{N}}$ such that $\text{dist}(S, T) \geq m$. Suppose $f : \overline{\mathbb{N}}^I \rightarrow \mathbb{R}$ and $g : \overline{\mathbb{N}}^J \rightarrow \mathbb{R}$ are bounded functions, where I and J are the cardinalities of S and T respectively. To complete the proof, it suffices to show that (see the sentence right

before Lemma 2.4)

$$|\text{Corr}(f(Y_k, k \in S), g(Y_k, k \in T))| \leq \varepsilon. \quad (3.13)$$

Now by Lemma 3.1, for each integer $n > \lambda$,

$$|\text{Corr}(f(Y_k^{(n)}, k \in S), g(Y_k^{(n)}, k \in T))| \leq \varepsilon. \quad (3.14)$$

If the left side of (3.13) is nonzero, then the left side of (3.14) converges to the left side of (3.13) as $n \rightarrow \infty$ by (3.12) and a routine calculation. Hence by (3.14), eq. (3.13) holds. That completes the proof. \square

4. PROOF OF THEOREM 1.3

As in the statement of Theorem 1.3, suppose $a \in (0, 1)$ and $\lambda > 0$. The argument here will be divided into four “steps”.

Step 1. *Construction of the sequence X .* For each integer ℓ (that is, each $\ell \in \mathbb{Z}$), let $Y^{(\ell)} := (Y_0^{(\ell)}, Y_1^{(\ell)}, Y_2^{(\ell)}, \dots)$ be a Markov chain with state space $\overline{\mathbb{N}}$, such that the distribution of this Markov chain $Y^{(\ell)}$ (on $\overline{\mathbb{N}}$) is the same as that of the Markov chain Y in Lemma 3.2. Let these Markov chains $Y^{(\ell)}$, $\ell \in \mathbb{Z}$ be constructed in such a way that they are independent of each other.

Just for convenient “bookkeeping” later on, for each $\ell \in \mathbb{Z}$ and each integer $k \leq -1$, define the degenerate random variable $Y_k^{(\ell)} \equiv 0$. For each $\ell \in \mathbb{Z}$, thereby extend the Markov chain $Y^{(\ell)}$ (retaining that notation) to the form $Y^{(\ell)} := (Y_k^{(\ell)}, k \in \mathbb{Z}) = (\dots, 0, 0, 0, Y_0^{(\ell)}, Y_1^{(\ell)}, Y_2^{(\ell)}, \dots)$. These random sequences $Y^{(\ell)}$, $\ell \in \mathbb{Z}$ are each a Markov chain, they are independent of each other, and they all have the same distribution (on, say, $\overline{\mathbb{N}}^{\mathbb{Z}}$). This extension does not change any of the dependence coefficients $\rho^*(Y^{(\ell)}, n)$.

Now for each $\ell \in \mathbb{Z}$ and each $j \in \overline{\mathbb{N}}$, the distribution of the random variable $Y_j^{(\ell)}$ is Poisson with mean λa^j (see Lemma 3.2(A)). Hence in particular, for each $\ell \in \mathbb{Z}$, $\sum_{j=0}^{\infty} EY_j^{(\ell-j)} < \infty$, and hence $\sum_{j=0}^{\infty} Y_j^{(\ell-j)} < \infty$ a.s. Define the sequence $X := (X_k, k \in \mathbb{Z})$ of the random variables as follows: For each $k \in \mathbb{Z}$,

$$X_k := \sum_{j=0}^{\infty} Y_j^{(k-j)} = \sum_{j=-\infty}^{\infty} Y_j^{(k-j)}. \quad (4.1)$$

Since the (nonstationary) Markov chains $Y^{(\ell)}$, $\ell \in \mathbb{Z}$ are independent of each other and have the same distribution, it follows from an elementary (if tedious) measure-theoretic argument that this random sequence X is strictly stationary. (Eq. (4.1) and the resulting stationarity of X are adapted from a scheme used in [13] to “convert” a nonstationary sequence to a stationary one preserving certain properties.)

Note that by (4.1) and the comments preceding it, one has (as in Lemma 2.5(A)) that for each $k \in \mathbb{Z}$,

$$\text{the distribution of } X_k \text{ is Poisson } (\lambda/(1-a)). \quad (4.2)$$

Step 2. *Verification of some features of the INAR(1) model with Poisson innovations.* For each integer k , referring to the comments preceding (4.1), define the

random variables U_k and V_k as follows:

$$U_k := \sum_{j=1}^{\infty} Y_j^{(k-j)} \text{ and } V_k := Y_0^{(k)}. \quad (4.3)$$

Then by (4.1), for each $k \in \mathbb{Z}$,

$$X_k = U_k + V_k. \quad (4.4)$$

By (4.3) and the comments preceding (4.1), one has that for each $k \in \mathbb{Z}$,

$$\text{the distribution of } V_k \text{ is Poisson } (\lambda). \quad (4.5)$$

By (4.1) and (4.3), for each $k \in \mathbb{Z}$,

$$\sigma(U_k) \subset \sigma(Y^{(\ell)}, \ell \leq k-1), \sigma(V_k) \subset \sigma(Y^{(k)}), \text{ and } \sigma(X_k) \subset \sigma(Y^{(\ell)}, \ell \leq k). \quad (4.6)$$

Since the Markov chains $Y^{(\ell)}$, $\ell \in \mathbb{Z}$ are independent of each other, one has by (4.6) that for each $k \in \mathbb{Z}$,

$$\sigma(V_k) \text{ is independent of } \sigma(U_j, V_j, X_j, j \leq k-1) \vee \sigma(U_k). \quad (4.7)$$

(Eqs. (4.4), (4.5), and (4.7) together have the interpretation that V_k is a ‘‘Poisson innovation’’.)

Next, suppose $k \in \mathbb{Z}$. Consider the independent random vectors

$$\left(Y_0^{(k-1)}, Y_1^{(k-1)} \right), \left(Y_1^{(k-2)}, Y_2^{(k-2)} \right), \left(Y_2^{(k-3)}, Y_3^{(k-3)} \right), \dots$$

By (4.1) and (4.3), the first coordinates of these random vectors add up to X_{k-1} , and the second coordinates add up to U_k . From the conditions in Lemma 3.2 (and the comments preceding (4.1)), the hypothesis of Lemma 2.5 is fulfilled.

Hence by Lemma 2.5(B), one has that for each $k \in \mathbb{Z}$ and each $x \in \overline{\mathbb{N}}$,

$$\begin{aligned} &\text{the conditional distribution of } U_k \text{ given } \{X_{k-1} = x\} \\ &\text{is binomial with parameters } x \text{ and } a. \end{aligned} \quad (4.8)$$

Also, from Lemma 2.5(C), one has that for each $k \in \mathbb{Z}$,

$$\left(\sigma(Y_j^{(k-1-j)}, j \geq 0), \sigma(X_{k-1}), \sigma(U_k) \right) \text{ is a Markov triplet.} \quad (4.9)$$

Step 3. *Two Markov triplets.* For each $\ell \in \mathbb{Z}$, define the σ -field

$$\mathcal{H}^{(\ell)} := \sigma(Y_{\ell-j}^{(j)}, j \in \mathbb{Z}). \quad (4.10)$$

By (4.1) and (4.3), for each $\ell \in \mathbb{Z}$,

$$\sigma(U_\ell, V_\ell, X_\ell) \subset \mathcal{H}^{(\ell)}. \quad (4.11)$$

Now for the rest of Step 3, let k be an arbitrary fixed integer. For this integer k , the task in the rest of Step 3 here is to establish two Markov triplets connected with the conditions in Definition 1.2.

For each $j \in \mathbb{Z}$, the ordered triplet of σ -fields

$$\left(\sigma(Y_u^{(j)}, u \leq k-2-j), \sigma(Y_{k-1-j}^{(j)}), \sigma(Y_{k-j}^{(j)}) \right)$$

is a Markov triplet. Since the Markov chains $Y^{(j)}$, $j \in \mathbb{Z}$ are independent, one has by (4.10) and a standard measure-theoretic argument that

$$\left(\bigvee_{i \leq k-2} \mathcal{H}^{(i)}, \mathcal{H}^{(k-1)}, \mathcal{H}^{(k)} \right)$$

is a Markov triplet. Hence by (4.11),

$$\left(\bigvee_{i \leq k-2} \mathcal{H}^{(i)}, \mathcal{H}^{(k-1)}, \sigma(U_k) \right) \quad (4.12)$$

is a Markov triplet.

Also, by (4.9) and the fact that $Y_j^{(k-1-j)} \equiv 0$ for $j \leq -1$,

$$\left(\mathcal{H}^{(k-1)}, \sigma(X_{k-1}), \sigma(U_k) \right) \quad (4.13)$$

is a Markov triplet.

Since $\sigma(X_{k-1}) \subset \mathcal{H}^{(k-1)}$ by (4.11), one has that for any event $C \in \sigma(U_k)$, by the sentences containing (4.12) and (4.13),

$$P \left(C \mid \bigvee_{i \leq k-1} \mathcal{H}^{(i)} \right) = P \left(C \mid \mathcal{H}^{(k-1)} \right) = P \left(C \mid \sigma(X_{k-1}) \right) \text{ a.s.};$$

and hence the ordered triplet

$$\left(\bigvee_{i \leq k-1} \mathcal{H}^{(i)}, \sigma(X_{k-1}), \sigma(U_k) \right)$$

is a Markov triplet. Hence by (4.11) again,

$$(\sigma(U_j, V_j, X_j, j \leq k-1), \sigma(X_{k-1}), \sigma(U_k)) \quad (4.14)$$

is a Markov triplet. Hence by (4.7) and a standard measure-theoretic argument,

$$(\sigma(U_j, V_j, X_j, j \leq k-1), \sigma(X_{k-1}), \sigma(U_k) \vee \sigma(V_k))$$

is a Markov triplet. Hence by (4.4),

$$(\sigma(X_j, j \leq k-1), \sigma(X_{k-1}), \sigma(X_k)) \quad (4.15)$$

is a Markov triplet.

Since $k \in \mathbb{Z}$ was arbitrary, the sequence X is by (4.15) a Markov chain, a property stipulated in Definition 1.2. Eq. (4.14) is (again for arbitrary $k \in \mathbb{Z}$) the other “Markov triplet” property stipulated in Definition 1.2. The other properties in Definition 1.2 (and the subsequent paragraph) were verified in (4.2), (4.4), (4.5), (4.7), and (4.8). That completes the verification that the sequence X is an INAR(1) model with Poisson innovations.

To complete the proof of Theorem 1.3, all that remains is to show that the sequence X is ρ^* -mixing.

Step 4. *Proof that X is ρ^* -mixing.* For each $j \in \mathbb{Z}$, define the “shifted random sequence” $\tilde{Y}^{(j)} := (\tilde{Y}_k^{(j)}, k \in \mathbb{Z})$ by $\tilde{Y}_k^{(j)} := Y_{k-j}^{(j)}$. Then by (4.1) (or (4.10)–(4.11)), for each $k \in \mathbb{Z}$, $\sigma(X_k) \subset \bigvee_{j \in \mathbb{Z}} \sigma(\tilde{Y}_k^{(j)})$. Hence by Lemma 2.1 and the first two paragraphs of Step 1, for any $n \in \mathbb{N}$,

$$\rho^*(X, n) \leq \sup_{j \in \mathbb{Z}} \rho^*(\tilde{Y}^{(j)}, n) = \sup_{j \in \mathbb{Z}} \rho^*(Y^{(j)}, n) = \rho^*(Y, n)$$

where the sequence Y is as in Lemma 3.2. By Lemma 3.2(B), that sequence Y is ρ^* -mixing. Hence X is ρ^* -mixing. That completes the proof of Theorem 1.3. \square

Acknowledgement. *The author thanks Mikhail Lifshits for his encouragement and helpful comments. The author also thanks the organizers of the Workshop on Recent Developments in Statistics for Complex Dependent Data, in Loccum, Germany, August 2015. The author was inspired by talks on INAR and related processes at that conference; and part of the research for this paper was done at that conference. That workshop was partly funded by the German Academic Exchange Service (DAAD) and by the Volkswagen Foundation (program: Niedersaechsisches Vorab for female professors of Lower Saxony).*

REFERENCES

- [1] R.C. Bradley. Every “lower psi-mixing” Markov chain is “interlaced rho-mixing”. *Stochastic Process. Appl.* 72 (1997) 221-239.
- [2] R.C. Bradley. A stationary rho-mixing Markov chain which is not “interlaced” rho-mixing. *J. Theor. Probab.* 14 (2001) 717-727.
- [3] R.C. Bradley. *Introduction to Strong Mixing Conditions*, Vol. 1. Kendrick Press, Heber City (Utah), 2007.
- [4] R.C. Bradley. On mixing properties of reversible Markov chains. *New Zealand J. Math.* (accepted for publication). arXiv:1403.4895v1 [math.PR] 19 Mar 2014.
- [5] R.C. Bradley and W. Bryc. Multilinear forms and measures of dependence between random variables. *J. Multivariate Anal.* 16 (1985) 335-367.
- [6] A.V. Bulinskii. On mixing conditions of random fields. *Theor. Probab. Appl.* 30 (1985) 219-220.
- [7] P. Csáki and J. Fischer. On the general notion of maximal correlation. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 8 (1963) 27-51.
- [8] J.G. Du and Y. Li. The integer-valued autoregressive (INAR(p)) model. *J. Time Series Anal.* 12 (1991) 129-142.
- [9] H.O. Hirschfeld. A connection between correlation and contingency. *Proc. Camb. Phil. Soc.* 31 (1935) 520-524.
- [10] I.A. Ibragimov. A note on the central limit theorem for dependent random variables. *Theor. Probab. Appl.* 20 (1975) 135-141.
- [11] A.N. Kolmogorov and Y.A. Rozanov. On strong mixing conditions for stationary Gaussian processes. *Theor. Probab. Appl.* 5 (1960) 204-208.
- [12] E. McKenzie. Some simple models for discrete variate time series. *Water Resour. Bull.* 21 (1985) 645-650.
- [13] R.A. Olshen. The coincidence of measure algebras under an exchangeable probability. *Z. Wahrsch. verw. Gebiete* 18 (1971) 153-158.
- [14] X. Pedeli and D. Karlis. Some properties of multivariate INAR(1) processes. *Comput. Statist. Data Anal.* 67 (2013) 213-225.
- [15] R. Peyre. Sharp equivalence between ρ - and τ -mixing coefficients. *Studia Math.* 216 (2013) 245-270.
- [16] M. Rosenblatt. *Markov Processes. Structure and Asymptotic Behavior*. Springer-Verlag, Berlin, 1971.
- [17] S. Schweer and C.H. Weiß. Compound Poisson INAR(1) processes: Stochastic properties and testing for overdispersion. *Comput. Statist. Data Anal.* 77 (2014) 267-284.
- [18] I. Silva and M.E. Silva. Parameter estimation for INAR processes based on high-order statistics. *REVSTAT* 7 (2009) 105-117.
- [19] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proceedings of the Sixth Berkeley Symposium on Probability and Statistics*, Vol. 2, 583-602. University of California Press, Los Angeles, 1972.
- [20] H.S. Witsenhausen. On sequences of pairs of dependent random variables. *SIAM J. Appl. Math.* 28 (1975) 100-113.

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY, BLOOMINGTON, INDIANA 47405, USA
E-mail address: `bradleyr@indiana.edu`